

A NOTE ON WHY ENFORCING DISCRETE MAXIMUM PRINCIPLES BY A SIMPLE A POSTERIORI CUTOFF IS A GOOD IDEA

CHRISTIAN KREUZER

ABSTRACT. Discrete maximum principles in the approximation of partial differential equations are crucial for the preservation of qualitative properties of physical models. In this work we enforce the discrete maximum principle by performing a simple cutoff. We show that for many problems this a posteriori procedure even improves the approximation in the natural energy norm. The results apply to many different kinds of approximations including conforming higher order and hp -finite elements. Moreover in the case of finite element approximations there is no geometrical restriction on the partition of the domain.

1. INTRODUCTION

Consider a function $u : \Omega \rightarrow \mathbb{R}$ on some bounded domain $\Omega \subset \mathbb{R}^d$ such that

$$(1.1) \quad -\Delta u + cu \leq 0$$

in the variational sense for some nonnegative c . Then for $u^+ = \max\{0, u\}$, we have the estimate

$$(1.2) \quad \sup_{\Omega} u \leq \sup_{\partial\Omega} u^+,$$

which is well-known as weak maximum principle; compare e.g. with [GT83].

Maximum principles usually reflect fundamental physical principles like e.g. the positivity of the density and it is desirable that numerical schemes respect physical principles. Therefore, for a conforming approximation U , computed by some numerical scheme, there arises the question if it satisfies a so called discrete maximum principle

$$(DMP) \quad \sup_{\Omega} U \leq \sup_{\partial\Omega} U^+.$$

Suppose for example that the approximation U is generated by a conforming finite element method. In this case there are plenty of results on discrete maximum principles; without attempting to provide a complete list, we refer to [CR73, RS82, BKK08, LH10, DKS12]. All those results have in common that they base on piece wise affine finite element spaces and on strong geometrical restrictions on the underlying partitions of the domain. To be more precise, they are basically restricted to non-obtuse or even acute simplicial meshes. Although there are some results on non-obtuse refinement in [KK11, KK05], it is clear that those restrictions introduce serious complications for the refinement and meshing of the domain Ω . The situation is even less satisfying for higher order finite elements let alone hp -finite elements. To our best knowledge for these schemes discrete maximum principles are known only in a relaxed sense [Sch80] or in very restrictive situations; see [HM81].

2010 *Mathematics Subject Classification.* 65N30 and 35J92 and 35J47 and 35J15.

Key words and phrases. discrete maximum principle and finite elements and nonlinear pde and p-Laplace and conforming approximation and reaction diffusion.

Based on an analysis of discrete Green's functions, Drăgănescu, Dupont, and Scott [DDS05] suggest that the discrete maximum principle for piece wise affine functions may only fail in a small region close to the boundary. This motivates to simply cutoff those small regions with unphysical behavior. In other words, we define

$$(1.3) \quad U^* = \min \left\{ U, \sup_{\partial\Omega} U^+ \right\} \quad \text{in } \Omega.$$

Obviously the modified function U^* satisfies the (DMP) independent of the polynomial degree or some underlying partition of Ω . Although so far this technique has lacked of mathematical justification, it is actually quite popular in engineering.

In this paper we overcome this drawback. In fact, using variational techniques from [DKS12, DLM98, BF02], we prove in section 2 that the modified function U^* is even a better approximation than U . To be more precise, if $u|_{\partial\Omega} = U|_{\partial\Omega}$, then we have

$$(1.4) \quad |||u - U^*||| \leq |||u - U|||$$

in the norm $|||\cdot|||^2 = \int_{\Omega} |\nabla \cdot|^2 + c|\cdot|^2 dx$, induced by the reaction diffusion differential operator in (1.1). In section 3 we generalize the techniques to systems of pde's and the p -Laplace operator.

We emphasize that the results are not restricted to finite element approximation nor to Galerkin schemes. In fact, the results can be applied to any conforming approximation of u .

2. ENFORCING THE DISCRETE MAXIMUM PRINCIPLE

In this section we shall prove the main result (1.3) for functions satisfying (1.1). To this end, for the bounded domain $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$, let $L^2(\Omega)$ and $H^1(\Omega)$ be the space of square integrable Lebesgue and Sobolev functions on Ω , respectively. We denote by $(H^1(\Omega))^*$ the dual of $H^1(\Omega)$ and by $H_0^1(\Omega)$ the subspace of functions in $H^1(\Omega)$ with vanishing trace on the boundary $\partial\Omega$.

The variational formulation of (1.1) reads as follows: Let $0 \leq c \in L^\infty(\Omega)$, i.e., a nonnegative essentially bounded function. We assume that $u \in H^1(\Omega)$ such that

$$(2.1) \quad \int_{\Omega} \nabla u \cdot \nabla v + c uv dx =: F(v) \leq 0 \quad \text{for all } v \in H_0^1(\Omega) \text{ with } v \geq 0 \text{ in } \Omega.$$

With this definition we have $F \in (H^1(\Omega))^*$ and it is well known that u satisfies the weak maximum principle (1.2); see [GT83]. Moreover, u is the unique minimizer of

$$(2.2) \quad \mathcal{J}(v) := \frac{1}{2} \int_{\Omega} |\nabla v|^2 + c|v|^2 dx - F(v)$$

in $u + H_0^1(\Omega)$. In other words, u is minimal among all functions in $H^1(\Omega)$ that coincide with u on the boundary $\partial\Omega$.

Let now $U \in H^1(\Omega)$ be some approximation to u . Note that there is no restriction on the kind of approximation beyond of that it is conforming; we will come back to this issue in Remark 2.4 and the Conclusion §4 below. It follows by standard arguments that

$$(2.3) \quad U^* := \min \left\{ U, \sup_{\partial\Omega} U^+ \right\} = (U - \sup_{\partial\Omega} U^+)^- + \sup_{\partial\Omega} U^+ \in H^1(\Omega).$$

Consequently, U^* satisfies the (DMP). If $U|_{\partial\Omega} = u|_{\partial\Omega}$, we have by (1.2) that

$$|u(x) - U^*(x)| = \begin{cases} |u(x) - U(x)|, & \text{if } U(x) \leq \sup_{\partial\Omega} U^+ \\ \sup_{\partial\Omega} U^+ - u(x), & \text{if } U(x) > \sup_{\partial\Omega} U^+ \end{cases} \leq |u(x) - U(x)|,$$

for almost every $x \in \Omega$. Hence, it follows

$$(2.4) \quad \|u - U^*\|_{L^2(\Omega)} \leq \|u - U\|_{L^2(\Omega)}.$$

Here $\|\cdot\|_{L^2(\Omega)}^2 := \int_{\Omega} |\cdot|^2 dx$ denotes the standard norm on $L^2(\Omega)$.

The corresponding estimate (1.4) for the energy norm is less obvious. To prove this estimate we need some more properties of the truncated function U^* . Similarly as in [DLM98] we have on the one hand, that

$$U^*(x) = \begin{cases} U(x) & \text{if } U(x) \leq \sup_{\partial\Omega} U^+, \\ \sup_{\partial\Omega} U^+ \geq 0, & \text{if } U(x) > \sup_{\partial\Omega} U^+, \end{cases} \quad x \in \Omega.$$

and hence we obtain

$$(2.5a) \quad U^* \leq U \quad \text{and} \quad |U^*| \leq |U| \quad \text{in } \Omega.$$

On the other hand we have

$$\nabla U^*(x) = \begin{cases} \nabla U(x), & \text{if } U(x) \leq \sup_{\partial\Omega} U^+, \\ 0, & \text{if } U(x) > \sup_{\partial\Omega} U^+, \end{cases} \quad x \in \Omega$$

and consequently

$$(2.5b) \quad |\nabla U^*| \leq |\nabla U| \quad \text{in } \Omega.$$

These observations are the key properties for proving the following result.

Theorem 2.1. *Suppose the conditions of this section. In particular, let $U \in H^1(\Omega)$ and let $U^* \in H^1(\Omega)$ be the modification of U according to (2.3). Then we have for the energy $\mathcal{J}: H^1(\Omega) \rightarrow \mathbb{R}$ defined in (2.2), that*

$$(2.6) \quad \mathcal{J}(U^*) \leq \mathcal{J}(U) \quad \text{and} \quad U^* - U \in H_0^1(\Omega).$$

Proof. The second claim is a direct consequence of the definition (2.3) of U^* . For the first claim we observe from (2.5b) that

$$\int_{\Omega} |\nabla U^*|^2 dx \leq \int_{\Omega} |\nabla U|^2 dx.$$

Furthermore, it follows from the second inequality in (2.5a) that

$$(2.7) \quad \int_{\Omega} |U^*|^2 dx \leq \int_{\Omega} |U|^2 dx$$

and since $F(v) \leq 0$ for all $v \in H^1(\Omega)$, $v \geq 0$, we obtain

$$F(U) - F(U^*) = F(U - U^*) \leq 0 \quad \Rightarrow \quad -F(U^*) \leq -F(U).$$

Combining these estimates together with the definition (2.2) of the energy proves the theorem. \square

Corollary 2.2. *Suppose the conditions of this sections and assume further that*

$$U - u \in H_0^1(\Omega), \quad \text{i.e., } u = U \text{ on } \partial\Omega.$$

Then we have

$$|||u - U^*||| \leq |||u - U|||$$

for the energy norm $|||\cdot|||^2 := \int_{\Omega} |\nabla \cdot|^2 + c|\cdot|^2 dx$ on $H_0^1(\Omega)$ induced by (2.1).

Proof. It follows from the assumption $u - U \in H_0^1(\Omega)$ that $U, U^* \in u + H_0^1(\Omega)$. We recall that u is the unique minimizer of the energy \mathcal{J} in $u + H_0^1(\Omega)$. Therefore, together with Theorem 2.1, we have

$$(2.8) \quad 0 \leq \mathcal{J}(U^*) - \mathcal{J}(u) \leq \mathcal{J}(U) - \mathcal{J}(u).$$

On the other hand, for arbitrary $v \in u + H_0^1(\Omega)$ we have $u - v \in H_0^1(\Omega)$ and it follows from (2.1) that

$$\begin{aligned} \mathcal{J}(v) - \mathcal{J}(u) &= \frac{1}{2} \int_{\Omega} |\nabla v|^2 + c|v|^2 \, dx - \frac{1}{2} \int_{\Omega} |\nabla u|^2 + c|v|^2 \, dx + F(u - v) \\ &= \frac{1}{2} \int_{\Omega} |\nabla v|^2 + c|v|^2 \, dx - \int_{\Omega} \nabla v \cdot \nabla u + c v u \, dx + \frac{1}{2} \int_{\Omega} |\nabla u|^2 + c|u|^2 \, dx \\ &= \frac{1}{2} \|u - v\|^2. \end{aligned}$$

Using this observation with $v = U$ respectively $v = U^*$ in (2.8) proves the claim. \square

Remark 2.3. For $c \equiv 0$ the maximum principle (1.2) reads as

$$\sup_{\Omega} u \leq \sup_{\partial\Omega} u.$$

Consequently, in this case, we define

$$U^* := \min \{U, \sup_{\partial\Omega} U\}.$$

Since $c \equiv 0$ we do not need to have the second estimate in (2.5a) in order to prove Theorem 2.1. All other estimates, namely the first estimate in (2.5a) and (2.5b), stay true for U^* as defined above. Therefore, Theorem 2.1 and Corollary 2.2 are still valid in this case; compare also with the examples of Section 3 below.

Remark 2.4. In the approximation of solutions to partial differential equations with finite elements one usually considers the error of the boundary values and the residual interior Ω , separately; see e.g. [BS08]. To be more precise, let u satisfy (2.1) and let G be a discrete approximation of u on $\partial\Omega$. We consider an approximation $U \in H_0^1(\Omega)$ with $U = G$ on $\partial\Omega$ to the weak solution $u_G \in H_0^1(\Omega)$ of the problem

$$-\Delta u_G + c u_G = -\Delta u + c u \quad \text{in } \Omega, \quad \text{and} \quad u_G = G \quad \text{on } \partial\Omega.$$

Then U satisfies Theorem 2.1 and Corollary 2.2 with u_G instead of u , i.e., let $U^* \in H^1(\Omega)$ be defined as in (2.3), then

$$\|u - U^*\| \leq \|u_G - U^*\| + \|u - u_G\| \leq \|u_G - U\| + \|u - u_G\|$$

The error $\|u - u_G\|$ can be estimated by means of the trace $u - G$ on $\partial\Omega$. Similar techniques can be applied in the case of a curved boundary.

3. EXTENSIONS

In the previous section, for the ease of presentation, we restricted ourselves to linear scalar valued reaction diffusion problems. However, the presented ideas can be generalized to more complicated problems. We shall present two examples.

3.1. Convex Hull Property. In this section we shall generalize the results of Section 2 to vector valued functions. In order to do so, we first have to generalize the cutoff process in (1.3) to higher dimensions. To this end, let $K \subset \mathbb{R}^m$, $m \in \mathbb{N}$, be a convex set and define $\Pi_K : \mathbb{R}^m \rightarrow K$ to be the closest point projection with respect to the Euclidean norm $|\cdot| : \mathbb{R}^m \rightarrow \mathbb{R}$. In other words

$$(3.1) \quad \Pi_K x := \operatorname{argmin}_{y \in K} |x - y|.$$

This definition can be extended to

$$\Pi_K : L^2(\Omega)^m \rightarrow L^2(\Omega)^m \quad \text{setting} \quad (\Pi_K \mathbf{v})(x) := \Pi_K \mathbf{v}(x).$$

It follows from the convexity of K , with elementary computations, that Π_K is 1-Lipschitz and hence a generalized chain rule implies $\Pi_K : H^1(\Omega)^m \rightarrow H^1(\Omega)^m$ with

$$(3.2) \quad |\nabla(\Pi_K \mathbf{v})(x)| \leq \operatorname{Lip}(\Pi_K) |\nabla \mathbf{v}(x)| = |\nabla \mathbf{v}(x)| \quad x \in \Omega;$$

compare with [BF02, AD90, AFP00]. This is the replacement for (2.5b).

Let $\mathbf{u} \in H^1(\Omega)^m$ be such that

$$\int_{\Omega} \nabla \mathbf{u} : \nabla \mathbf{v} \, dx = 0 \quad \text{for all } \mathbf{v} \in H_0^1(\Omega)^m.$$

It is well known that \mathbf{u} satisfies the convex hull property

$$\mathbf{u}(\Omega) \subset \operatorname{conv hull} \mathbf{u}(\partial\Omega)$$

(see [BF02]) and that \mathbf{u} is the unique minimizer of the energy

$$\mathcal{J}(\mathbf{v}) := \int_{\Omega} \frac{1}{2} |\nabla \mathbf{v}|^2 \, dx \quad \text{in } \mathbf{u} + H_0^1(\Omega)^m.$$

Let $\mathbf{U} \in H^1(\Omega)^m$ be some approximation of \mathbf{u} and define

$$\mathbf{U}^* := \Pi_K \mathbf{U}, \quad \text{with } K := \operatorname{conv hull} \mathbf{U}(\partial\Omega).$$

Consequently \mathbf{U}^* satisfies the discrete convex hull property

$$\mathbf{U}^*(\Omega) \subset \operatorname{conv hull} \mathbf{U}^*(\partial\Omega)$$

and it follows from (3.2) similar as in the proof of Theorem 2.1 that $\mathbf{U}^* - \mathbf{U} \in H_0^1(\Omega)^m$ and

$$(3.3) \quad \mathcal{J}(\mathbf{U}^*) = \frac{1}{2} \int_{\Omega} |\nabla \mathbf{U}^*|^2 \, dx \leq \frac{1}{2} \int_{\Omega} |\nabla \mathbf{U}|^2 \, dx = \mathcal{J}(\mathbf{U}).$$

Moreover, as in Corollary 2.2, we obtain for $\|\cdot\|^2 = \int_{\Omega} |\nabla \cdot|^2 \, dx$ that

$$(3.4) \quad \|\mathbf{u} - \mathbf{U}^*\| \leq \|\mathbf{u} - \mathbf{U}\| \quad \text{if } \mathbf{u} - \mathbf{U} \in H_0^1(\Omega)^m.$$

Remark 3.1. *The results of this section can be generalized to reaction diffusion problems. To this end we observe that, in order to prove (3.3), we need a replacement of (2.5a) for vector-valued functions. Note that for reaction diffusion problems the convex hull property reads as*

$$\mathbf{u}(\Omega) \subset \operatorname{conv hull} (\{0\} \cup \mathbf{u}(\partial\Omega)).$$

Therefore, defining $K := \operatorname{conv hull}(\{0\} \cup \mathbf{U}(\partial\Omega))$ we obtain $|\mathbf{U}^(x)| \leq |\mathbf{U}(x)|$, $x \in \Omega$. This is the required replacement of (2.5a).*

3.2. Nonlinear Problems. Recalling the proof of Theorem 2.1 and (2.5), we observe that formally (2.6) holds for energies

$$\mathcal{J}(v) = \int_{\Omega} \mathcal{F}(x, |v|, |\nabla v|) dx - F(v)$$

with $\mathcal{F} : \Omega \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ being monotone in its second and third variable. In particular, Theorem 2.1 can be applied to many nonlinear problems.

As an example, we consider the nonlinear p -Laplace problem. To this end, for fixed $p \in (1, \infty)$, $\frac{1}{p} + \frac{1}{q} = 1$, let $W^{1,p}(\Omega)$ be the space of p integrable functions on Ω with p -integrable weak derivatives. Let $u \in W^{1,p}(\Omega)$ such that

$$-\operatorname{div} |\nabla u|^{p-2} \nabla u =: F \leq 0 \quad \text{in } \Omega;$$

in the distributional sense. Then u is the unique minimizer of the energy

$$\mathcal{J}(v) := \int_{\Omega} \frac{1}{p} |\nabla v|^p dx - F(v) \quad \text{in } u + W_0^{1,p}(\Omega).$$

Here $W_0^{1,p}(\Omega)$ is the subspace of functions in $W^{1,p}(\Omega)$ with vanishing trace on $\partial\Omega$. Moreover, it is well known that u satisfies

$$\sup_{\Omega} u \leq \sup_{\partial\Omega} u;$$

see e.g. [DLM98]. Let $U \in W^{1,p}(\Omega)$ be some approximation to u and let

$$U^* := \min \{U, \sup_{\partial\Omega} U\}.$$

This implies (2.5b). Note that the differential operator contains no reactive term and thus we do not require (2.5a) to conclude, similar as in the proof of Theorem 2.1, that

$$\mathcal{J}(U^*) \leq \mathcal{J}(U) \quad \text{and thus} \quad \mathcal{J}(U^*) - \mathcal{J}(u) \leq \mathcal{J}(U) - \mathcal{J}(u).$$

In order to prove a result analog to Corollary 2.2 it remains to correlate the energy difference to a reasonable measure of distance. This can be done using the so-called quasi-norm introduced by Barrett and Liu in [BL93]. It follows from [Kre12, Lemma 13] that there exist constants $C, c > 0$ such that

$$c(\mathcal{J}(v) - \mathcal{J}(u)) \leq \|v - u\|_{(\nabla u)}^2 := \int_{\Omega} (|u| + |v|)^{p-2} |u - v|^2 dx \leq C(\mathcal{J}(v) - \mathcal{J}(u))$$

for all $v \in u + W_0^{1,p}(\Omega)$. Hence, with $v = U$ respectively $v = U^*$, it follows from above, that

$$\sqrt{c} \|u - U^*\|_{(\nabla u)} \leq \sqrt{C} \|u - U\|_{(\nabla u)}.$$

4. CONCLUSION

Maximum principles often reflect physical behavior of solutions and it is therefore desirable that numerical approximations satisfy a maximum principle as well. In this paper we enforce the maximum principle by performing a simple cutoff to the approximation. We show that for many problems this truncation even improves the approximation. Hence, all error estimates and convergence results for the approximation can directly be applied to the error of the truncated function.

We emphasize that we do not specify the sort of approximation beyond that it is conforming. Therefore, among others, all presented results apply to all conforming finite element approximation including high order elements and conforming hp -methods. There is no restriction on the underlying partition of Ω like non obtuseness of the triangulation and the results even apply to partitions involving complicated element geometries.

REFERENCES

- [AD90] L. Ambrosio and G. DalMaso, *A general chain rule for distributional derivatives*, Proc. Amer. Math. Soc. **108** (1990), 691–702.
- [AFP00] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 2000.
- [BF02] M. Bildhauer and M. Fuchs, *Partial regularity for a class of anisotropic variational integrals with convex hull property*, Asymptot. Anal. **32** (2002), no. 3-4, 293–315.
- [BKK08] J. H. Brandts, S. Korotov, and M. Křížek, *The discrete maximum principle for linear simplicial finite element approximations of a reaction-diffusion problem*, Linear Algebra Appl. **429** (2008), no. 10, 2344–2357.
- [BL93] J. W. Barrett and W. B. Liu, *Finite element approximation of the p -Laplacian.*, Math. Comput. **61** (1993), no. 204, 523–537.
- [BS08] S. Brenner and R. Scott, *The mathematical theory of finite element methods*, Springer Texts in Applied Mathematics 15, 2008 (English).
- [CR73] P. G. Ciarlet and P.-A. Raviart, *Maximum principle and uniform convergence for the finite element method*, Comput. Methods Appl. Mech. Engrg. **2** (1973), 17–31.
- [DDS05] A. Drăgănescu, T. F. Dupont, and L. R. Scott, *Failure of the discrete maximum principle for an elliptic finite element problem*, Math. Comp. **74** (2005), no. 249, 1–23 (electronic).
- [DKS12] L. Diening, C. Kreuzer, and S. Schwarzacher, *Convex hull property and maximum principles for finite element minimizers of general convex functionals*, Numerical Analysis Technical Report No. 1501, University of Oxford, submitted, 2012.
- [DLM98] A. D’Ottavio, F. Leonetti, and C. Musciano, *Maximum principle for vector-valued mappings minimizing variational integrals*, Atti Sem. Mat. Fis. Univ. Modena **46** (1998), no. suppl., 677–683, Dedicated to Prof. C. Vinti (Italian) (Perugia, 1996).
- [GT83] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Springer, 1983 (English).
- [HM81] W. Höhn and H.-D. Mittelmann, *Some remarks on the discrete maximum-principle for finite elements of higher order*, Computing **27** (1981), no. 2, 145–154.
- [KK05] S. Korotov and M. Křížek, *Global and local refinement techniques yielding nonobtuse tetrahedral partitions*, Comput. Math. Appl. **50** (2005), no. 7, 1105–1113.
- [KK11] ———, *Nonobtuse local tetrahedral refinements towards a polygonal face/interface*, Appl. Math. Lett. **24** (2011), no. 6, 817–821.
- [Kre12] C. Kreuzer, *Analysis of an adaptive Uzawa finite element method for the nonlinear Stokes problem*, Math. Comp. **81** (2012), 21–55.
- [LH10] X. Li and W. Huang, *An anisotropic mesh adaptation method for the finite element solution of heterogeneous anisotropic diffusion problems*, J. Comput. Phys. **229** (2010), no. 21, 8072–8094.
- [RS82] V. Ruas Santos, *On the strong maximum principle for some piecewise linear finite element approximate problems of nonpositive type*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **29** (1982), no. 2, 473–491.
- [Sch80] A. H. Schatz, *A weak discrete maximum principle and stability of the finite element method in L_∞ on plane polygonal domains. I*, Math. Comp. **34** (1980), no. 149, 77–91.

CHRISTIAN KREUZER, FAKULTÄT FÜR MATHEMATIK, RUHR-UNIVERSITÄT BOCHUM, UNIVERSITÄTSTRASSE 150, D-44801 BOCHUM, GERMANY

URL: <http://www.ruhr-uni-bochum.de/ffm/Lehrstuehle/Kreuzer/>

E-mail address: christian.kreuzer@rub.de